

# Step Towards to Improve the Voluntary Interruption of Pregnancy by Means of Business Intelligence

## ABSTRACT

*With the implementation of Information and Communication Technologies in the health sector, it became possible the existence of an electronic record of information for patients, enabling the storage and the availability of their information in databases. However, without the implementation of a Business Intelligence (BI) system, this information has no value. Thus, the major motivation of this paper is to create a decision support system that allows the transformation of information into knowledge, giving usability to the stored data. The particular case addressed in this chapter is the Centro Materno Infantil do Norte, in particular the Voluntary Interruption of Pregnancy unit. With the creation of a BI system for this module, it is possible to design an interoperable, pervasive and real-time platform to support the decision-making process of health professionals, based on cases that occurred. Furthermore, this platform enables the automation of the process for obtaining key performance indicators that are presented annually by this health institution. In this chapter, the BI system implemented in the VIP unity in CMIN, some of the KPIs evaluated as well as the benefits of this implementation are presented.*

**Keywords:** Business Intelligence, Voluntary Interruption of Pregnancy, Data Mining, Centro Materno Infantil Norte.

## INTRODUCTION

The use of Information and Communication Technologies (ICT) are increasingly, occupying an important place in society. The health sector is no exception, as these among other things, can provide complete and reliable information for healthcare professionals, allowing to support their clinical and administrative decisions and consequently decreasing medical errors associated to these decisions (Pinto, 2009).

Besides, ICT have a high potential to facilitate information sharing, communication and collaboration between health professionals, increasing the quality and efficiency of the health system as well as the use of Electronic Health Records (EHR) (Abelha et al., 2007; Khodambashi, 2013; Portela et al., 2010) is one of the goals.

In recent times, Business Intelligence (BI) technologies have been the target of interest to health professionals and to the Information Technology (IT) professionals, due to its applicability in EHR (Bonney, 2013). BI is a process that encompasses several methodologies, applications and technologies for collecting, storing, manipulating, analysing and providing access to data in order to help enterprise users making better and faster business decisions. Thus BI has the ability to operationalize the repository content of EHR in supporting evidence-based practice and improving the quality of healthcare delivery (Bonney, 2013; Mettler & Vimarlund, 2009; Portela et al., 2010).

In the case of healthcare organizations, the majority of clinical data documenting their daily activities are stored in a Relational Database Management System (RDBMS). Because of the extensive amount of information, this information is stored in different ways and therefore highly heterogeneous with each other. On the other hand, a decision-making process, where it is necessary to integrate multiple data

provided by clinical, medical, financial and administrative systems and where the sources are quite heterogeneous, large and complex becomes extremely important to meet the data quality that directly interferes in the success of the Knowledge Discovery Database (KDD) process (Mettler & Vimarlund, 2009; Raquel & Oliveira, 2012). So, with this increasing amount of information there is also a corresponding need to apply Data Mining (DM) technologies to extract knowledge from information / data stored in databases and providing real-time decisions (Bonney, 2013). Most clinical data are not structured and the DM techniques work well with structured data. It is inferred another advantage to using BI as a decision support technology since it allows the combination of structured and unstructured data (Bonney, 2013).

Furthermore, it can be stated that the Information Systems interoperability in an institution is one of the key factors in the decision-making process. Interoperability (Cardoso, Marins, Portela, Santos, et al., 2014) ensures systems standardization and allows passing all the inherent complexity of the different data sources ensuring data quality.

Thus, this chapter is focused on presenting the architecture of a global BI platform that can be used in Maternity Care Institution, taking as a case study the Centro Materno Infantil do Norte (CMIN) in Porto, Portugal. It was planned apply BI and DM techniques in order to develop a platform where it is possible generate useful Key Performance Indicators (KPIs) (clinical and management) for healthcare professionals in the context of the Voluntary Interruption of Pregnancy (VIP) module. For the development of this platform, one Data Warehouse (DW) was built using the star schema and a set of DM models were induced. For better accuracy of the solution and validity of information generated, all the work has been pleading with clinicians and specialists.

The solution in development is a pervasive real-time web-based BI application and due to their characteristics allows it to be accessed anywhere and anytime.

Besides the introduction, this chapter includes eight sections. The second section is related to the background knowledge, where a brief look is taken about Maternity Care and Interoperability and it is described the VIP process. Subsequently, section three describes the process of Knowledge Discovery in Databases and section four outlines the BI system that is being implemented in CMIN and some of the indicators to get at the end. Finally, some discussions and conclusions were made and future work was suggested.

## **BACKGROUND AND RELATED WORK**

### **Maternity Care and Voluntary Interruption of Pregnancy**

The Centro Hospitalar do Porto (CHP) results from the union of three hospitals, namely, the Hospital de Santo António (HSA), Hospital Joaquim Urbano (HJU) and Centro Materno Infantil do Norte (CMIN). Each one of these units were separate entities until the creation of the CHP in 2007. With the creation of CHP, women who need urgent care in the specialties of Gynaecological and Obstetrics are directed to MJD, now CMIN, which was not previously checked (Cabral et al., 2013)

CMIN is a health institution, which one of the modules is VIP. In this module, clinically non-surgical methods are used for this purpose. More specifically, World Health Organization (WHO) recommends a drug method that has been proven safe and it is used quite effectively. This method consists of the combination of medication (*mifepristone* and *misoprostol*).

Furthermore, VIP is a procedure, which involves several steps. The first phase consists of a physician appointment followed of a period of three days, where the patient needs to consider its decision. Then, the patient is administered with a dosage of medication, performed at ambulatory in CMIN. A triage is also performed by the nursing staff in order to verify if the patient is capable to make the administration of the second medication dosage at home or if she needs to be monitored by the nursing team in CMIN.

After the medication administration phase, a new medical consultation is performed, where the patient is examined in order to determine whether the procedure was successful. If the opposite occurs, the

Interruption of Pregnancy (IoP) was not achieved or the procedure was incomplete (resulting in the patient admission in CMIN) (Valente, Cristina, Rosário, & Alcina, 2012).

In this last phase, two situations may occur: the patient is consulted and he is examined by the physician or the patient does not attend to the consultation. In the first case, the patient is evaluated and it is reported if IoP was successful or not. On the other hand, if ovular remains was verified by ultrasound (incomplete IOP), it is necessary the patient hospitalization, and if IOP was not achieved, it is necessary to repeat the process.

In the second case, the patient was not evaluated, being their condition and the result of the procedure unknown. This situation characterizes the risk patients and can originates some problems, in particular, health risks associated to the patient and their fetus due a lack of medical supervision.

The idea of implementing a BI platform is to provide a follow up able to facilitate the health professional's work and improve the quality of care as is for example the triage system (Abelha et al., 2015). They were responsible for obtaining reports, which should contain indicators related to the information recorded during the proceedings. Nowadays these reports are held annually in paper and are used for evaluation of CMIN, as is the example of VIP module (Valente et al., 2012).

## Interoperability System

At Centro Hospitalar do Porto (CHP), the interoperability between information systems is ensured by Agency for Integration, Archive and Diffusion of Medical Information (AIDA). This system of intelligent agents (Cardoso, Marins, Portela, Abelha, & Machado, 2014) allows communication of the different CHP systems, as well as standardization of clinical systems, overcoming the medical and administrative complexity of the different information sources (Marins et al., 2014). Thus, AIDA ensures Information sending and receiving from the hospital and is able to managing, storing and responding to the requests for information made from third parties (Machado, Alves, Abelha, & Neves, 2007; Peixoto, Santos, Abelha, & Machado, 2012). All information systems of CHP are connected through AIDA, for example, the SAPE (Nursing Support System) and SONHO (Clinical Information System) (Machado et al., 2007; Peixoto et al., 2012).

## KNOWLEDGE DISCOVERY IN DATABASES

This section is focused on Knowledge Discovery in Databases (KDD) and DM. In this section all the procedures performed by CRISP-DM methodology, to obtain the DM models associated with VIP unit in CMIN are described.

The Knowledge Discovery in Database (KDD) process is a set of ongoing activities that share the knowledge discovered from databases. According to Fayyad et al. (Fayyad, Piatetsky-shapiro, & Smyth, 1996), KDD process consists in five stages, which were followed in this work:

- **Selection:** Selection of the data needed to perform the Data Mining (DM).
- **Pre-processing:** This step included cleaning and processing of data in order to make them consistent.
- **Transformation:** In this phase the data were worked out according to the target.
- **Data Mining:** At this stage, the objectives to be achieved and the type of result wanted to achieve were defined. According to the type of desired result, it was defined the type of task being performed (classification, segmentation, summarization, dependency modeling) and identified the technique to be used (decision trees, association rules, linear regression, neuronal networks, among others). Subsequently, it was applied the selected data mining technique to the data set to obtain patterns.
- **Interpretation/Evaluation:** Consisted in the interpretation and evaluation of the patterns obtained. The validity of the results obtained are was verified by applying the patterns found at new dataset (Palaniappan & Awang, 2008).

The KDD process refers to the whole process of discovering useful knowledge in data, while DM refers to the application of algorithms to extract data models. Until 1995, KDD and DM terms were considered interchangeably. Now DM represents a phase of the Knowledge Discovery in Databases process (KDD) and consists in finding patterns or relationships that may exist in the data stored in data repositories.

The knowledge to be discovered must satisfy three properties: it must be correct (as much as possible), should be understandable by human users, and should be interesting useful/new. Still, the method of knowledge discovery must have the following three characteristics: it must be effective (accurate), generic (applicable to various data types) and flexible (easily modifiable) (Steiner, Soma, & Shimizu, 2006).

Within the VIP unit in CMIN, two DM problems were addressed. The first problem was related to the capability of predicting the most appropriate place for the administration of the second drug dosage (Brandão et al., 2014). With the second problem, it was intended to evaluate the VIP process and consequently identify whether the patient will belong to a risk group - probability of a patient not attend to the revision consultation (Machado et al., 2015).

The methodology addressed for both problems was the Cross Industry Standard Process for Data Mining (CRISP-DM). CRISP-DM divides the DM process into six major phases, as can be seen in Figure 1 (Chapman et al., 2000; Cios, Pedrycz, Swiniarski, & Kurgan, 2007; Gonçalves et al., 2013).



Figure 1. The CRISP-DM process. Source: (Chapman et al., 2000)

The steps followed were:

- **Business Understanding:** this initial phase of the project DM focused on understanding the objective of the project from a business perspective, defining a preliminary plan to achieve the goal. In this phase, the DM problems can be translated in two questions, namely, "What is the probability of a pregnant woman carry out the second drug phase with monitoring of nursing staff?" and "What is the probability of a patient belonging to the risk group of patients?".
- **Data Understanding:** comprised data collection and startup activities for data comprehension, identifying problems or interesting sets. In this phase, for each problem it was extracted data from SAPE and AIDA and it was analyzed the quality of the possible variables to be used in the process. The sample covers the period between 01.01.2012 and 31.12.2012. It corresponds to 1124 VIP cases. Furthermore, a statistical analysis was performed, showing that the data had quality (low

number of noise values), however the data needed to be prepared in order to be incorporated into the DM models.

- **Data Preparation:** in the preparation phase, all the tasks involved in creating cases that were used to build the models were included. Data preparation tasks were likely to be performed multiple times, in no particular order. Tasks included building the table of cases, selection of attributes, data cleaning and transformation. Additionally, it was possible add new attributes calculated based on existing ones. The preparation phase of data can significantly improve the information that can be discovered through DM. In these particular cases, the variables that suited the problems were selected. Subsequently, the selected data were subjected to a pre-processing phase where all records containing unfilled or noise fields were eliminated. Some of the procedures performed to eliminate noise from the data were replacing the comma by point to separate the decimals on numerical variables and also deleting text associated to these variables. After those transformations of the data, it became possible to build the table of cases with some of the scenarios considered most suitable.
- **Modeling:** several modeling techniques were applied. Thus, it was common return to data preparation during this phase. In this step, for each problem was built a table of cases composed by 10 scenarios that yielded the best results, arising from different combinations of variables, as can be seen in Tables 1 and 2. In each scenario, the target variable was presented, as well as other variables considered crucial to the creation of forecast models. After set the table of cases, the data were submitted to DM techniques, selected in order to be able to identify the best forecasting model for this DM problem. In this case, the DM techniques used were Generalized Linear Model, Support Vector Machine, Naïve Bayes and Decision Tree.

The variables used in the first problem were: LASD (Local administration of the second dose), Age, N\_VIP (number of previous VIP), Gesta (number of previous pregnancies), Para (number of births), PS (Professional Status), CM (Contraceptive Method), WG (Weeks of Gestation), AC\_1 (Age Class 1), AC\_2 (Age Class 2), WGC (Weeks of Gestation Class)

*Table 1. Representation of the variables used in each of the models in the first DM problem.*

|             | LASD | Age | N_VIP | Gesta | Para | PS | CM | WG | AC_1 | AC_2 | WGC |
|-------------|------|-----|-------|-------|------|----|----|----|------|------|-----|
| Scenario 1  | X    | X   | X     | X     | X    | X  | -  | -  | -    | -    | -   |
| Scenario 2  | X    | X   | -     | X     | X    | -  | -  | -  | -    | -    | -   |
| Scenario 3  | X    | -   | -     | X     | X    | -  | -  | -  | -    | -    | -   |
| Scenario 4  | X    | X   | X     | X     | X    | -  | -  | -  | -    | -    | -   |
| Scenario 5  | X    | -   | X     | X     | X    | X  | -  | -  | X    | -    | -   |
| Scenario 6  | X    | -   | X     | X     | X    | X  | -  | -  | -    | X    | -   |
| Scenario 7  | X    | -   | X     | X     | X    | X  | X  | -  | -    | X    | -   |
| Scenario 8  | X    | -   | X     | X     | X    | X  | -  | X  | -    | X    | -   |
| Scenario 9  | X    | -   | X     | X     | X    | X  | -  | -  | -    | X    | X   |
| Scenario 10 | X    | -   | X     | X     | X    | X  | X  | -  | -    | X    | X   |

In the second problem other variables were used: RC (Revision Consultation), Age, N\_VIP (number of previous VIP), Gesta (number of previous pregnancies), Para (number of births), PS (Professional Status), CM (Contraceptive Method), WG (Weeks of Gestation)

Table 2. Representation of the variables used in each of the models in the second DM problem.

|             | <i>RC</i> | <i>Age</i> | <i>N_VIP</i> | <i>Gesta</i> | <i>Para</i> | <i>PS</i> | <i>CM</i> | <i>WG</i> |
|-------------|-----------|------------|--------------|--------------|-------------|-----------|-----------|-----------|
| Scenario 1  | X         | X          | X            | X            | X           | X         | X         | X         |
| Scenario 2  | X         | X          | X            | X            | X           | X         | X         | -         |
| Scenario 3  | X         | X          | X            | X            | X           | X         | -         | -         |
| Scenario 4  | X         | X          | X            | X            | X           | -         | -         | -         |
| Scenario 5  | X         | X          | X            | -            | -           | -         | -         | -         |
| Scenario 6  | X         | X          | X            | X            | X           | X         | -         | X         |
| Scenario 7  | X         | X          | X            | X            | X           | -         | -         | X         |
| Scenario 8  | X         | X          | X            | X            | X           | -         | X         | X         |
| Scenario 9  | X         | X          | X            | X            | X           | -         | X         | -         |
| Scenario 10 | X         | X          | X            | -            | -           | -         | X         | X         |

- **Evaluation:** there was built models appearing to have great quality from a data analysis perspective. However, it was necessary to check whether the model met the business goals. To evaluate the results achieved by the DM models, statistical metrics were used. In these metrics tree parameters were evaluated: specificity, sensitivity and accuracy.
- **Deployment:** the knowledge gained by the model was organized and presented in a way that the health profession can use. In this particular case, the knowledge gained was organized and presented so that the nurses can use them. These DM processes were integrated in the BI platform, implemented in CMIN.

The next section describes all the steps carried out during the project for the design and implementation of BI platform.

## BUSINESS INTELLIGENCE

Currently, a BI platform can be defined as "a set of tools, programmed and integrated technologies and products that are used for the collection, analysis and presentation of data." The BI has the ability to turn the stored information into useful knowledge and provide the right user, in time, to support the decision-making process (Bonney, 2013; Reinschmidt & Francoise, 2000).

With the increase of electronic records and with increasing pressure to have faster and more effective decisions, which are often responsible for the occurrence of medical malpractice, it becomes appropriate to construct a support system to the specific decision making process for the VIP unit in CMIN, using BI technology. Moreover, nursing staff performs annual reports in paper, containing indicators related to the recorded information during the VIP procedures. The work performed by nurses for the creation of the indicators is no longer necessary, by the use BI technologies, the entire process is automatized and executed in real-time.

In the development of BI platform for the VIP unit in CMIN (Brandão et al.; Pereira et al.), it was followed the Kimball approach (El-Sappagh, Hendawi, & El Bastawissy, 2011; Kimball & Ross, 2002).

This method identifies a sequence of tasks and highlights activities that must be performed in parallel, entailing the following steps, depicted in Figure 2:

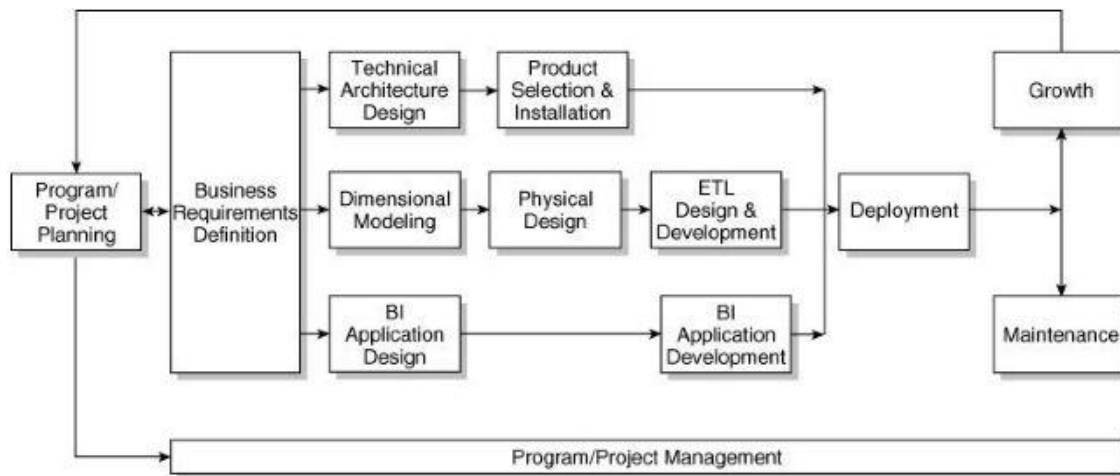


Figure 2. Kimball's Approach Diagram, Source (Kimball & Ross, 2002)

- **Program/Project Planning:** This phase includes project planning, understanding the requirements of organization and the identification of tasks. Thus, at this stage, some tasks were performed: researching important indicators for the IVG unit of CMIN, an initial planning of all stages of the Kimball approach and setting deadlines for these same steps defined.
- **Program/Project Management:** This stage is monitored throughout the project from the initial stage to the end, responsible for strengthening the project plan initially established.
- **Business Requirements Definition:** At this stage, it was performed planning / research all indicators required in the VIP unit with the help of health professionals and professionals of Information Technology (IT) of CMIN in order to obtain a more precise idea of the end requirements of the BI platform.

In the diagram of Figure 2, it can be seen that there are three sets of project activities occurring simultaneously. These sets are associated to the technological part, from the data acquisition until the development and deployment of BI application. The arrows in the diagram indicate the flow of activities within each of parallel paths. Dependencies between tasks are illustrated by vertical alignment of the boxes of the respective task (Kimball & Ross, 2002). These routes are:

- **Technological tasks**, including:
  - **Technical Architecture Design:** A technical scheme of the overall architecture is realized.
  - **Product Selection and Installation:** In this step must be selected and installed all of the appropriate Business Intelligence tools in order to be able to build the BI solution. At this stage, some tools were installed and tested, including **Oracle Database**, as applied to the database; **Oracle SQL Developer** as a tool for data extraction and implementation of ETL procedures; and **Pentaho BI Suite** as a tool for OLAP analysis and the creation of dashboards. This last selected tool for the development of BI application was chosen based on a comprehensive study on open-source BI tools.
- **Data tasks**, including:

- **Dimensional Modeling:** The dimensional model is a DW projection technique, which aims to provide a facilitated support to end users of the consultations. The dimensional model is constituted by the facts table and by the dimensions tables. The facts are typically numerical values, which can be aggregated, while the dimensions are hierarchies and values groups that define events. For the projection of dimensional model can be used various types of architectures. In structuring the DW dimensional model implemented in CMIN, the star schema was used because it was the most appropriate for the intended structure for the provision of data.
- **Physical Design:** In this phase was defined the data physical structure involving tasks such as setting the database environment and appropriate security setting.
- **ETL Design and Development:** This phase included extraction, transformation and loading of the data. The extraction consists of two stages, the initial extraction and changed data extraction. In the first stage, extracting the initial data occurs before DW loading, while the second stage is a continuous update of the DW, such as continuous extraction of the inherent data from CMIN databases. The processing corresponds to cleaning, integration and optimization of the data in order to achieve uniform, consistent and accurate data. Finally, there is the loading phase, wherein the data subjected to the above steps were inserted in a multidimensional structure (DW) targeted to a specific type of end users and applications systems. To carry out this process, procedures and functions have been implemented through SQL Developer. In the ETL procedures, some of the goals were achieved by the elimination of null values, the elimination of redundant data and the processing of information. In addition, procedures were used to standardize the information, such as correcting spelling errors, replacing the comma for the point at values with decimal places, among others.
- **The Practical Application of BI, includes:**
  - **BI Application Design:** At this stage it was necessary to identify the candidates' BI applications and appropriate navigation interfaces to meet the needs of CMIN professionals.
  - **BI Application Development:** At this stage, some tasks were performed, such as construction and validation of specific operational BI applications and navigation portal. Thus, it was used the Pentaho BI Suite for building OLAP cubes and visualization of performance indicators required by the CMIN.
- **Deployment:** In this phase the development of BI application was started. A proper planning of the work was crucial. It was ensured that the paths of technology, data and BI application were tested and they were working correctly together.
- **Maintenance:** Occurs when the system is in production. Includes technical operational tasks that are necessary to maintain the system with optimal performance (monitoring the use, performance tuning, index maintenance, system back-up).
- **Growth:** DW systems tend to expand, if they are successful. In this case, it must be restarted again Kimball cycle and to develop the issues already established, focusing on new requirements. In the case of CMIN, the BI platform can be expanded with new modules of interest in the organization to support the decision-making process (El-Sappagh et al., 2011; Kimball & Ross, 2002).

In the next section, the BI platform developed for the VIP unit of CMIN is presented, where are incorporated the indicators and the DM models developed.



## BUSINESS INTELLIGENCE PLATFORM FOR VIP

The BI platform for the VIP unit in CMIN is based on the Web following pervasive characteristics (Filipe Portela, Jorge Aguiar, Manuel Filipe Santos, Álvaro Silva, & Fernando Rua, 2013; C. F. Portela, Santos, Silva, Machado, & Abelha, 2011; F. Portela et al., 2011; F. Portela et al., 2013; F. Portela, Santos, & Vilas-Boas, 2012), which can be accessed anytime and anywhere, according to the access privileges. In this application are represented all the required indicators and also other indicators that can be created by users through the OLAP technology. This platform allows a more detailed analysis at the date level, as it could be viewing the level of the year, the months and days. Some of these indicators are:

- **Characterization of the patient group by number of pregnancies and date:** identifies the number and percentage of patient by number of previously pregnancies in a selected time period.
- **Characterization of the patient group by number of children and by date:** identifies the number and percentage of patient by number of children they have, in a selected time period.
- **Characterization of the patient group by number of previous VIP experiences and by date:** identifies the number and percentage of patient who have previous experience of VIP in a selected time period.
- **Characterization of the patient group concerning the revision consultation for date:** displays the number and proportion of patient who were present at the revision consultation, corresponding to the appointment made at the end of VIP process for a certain time interval.
- **Characterization of the patient group** based on the final outcome of the VIP process and by date: identifies the number of patients and the percentage of the three types of possible end results in the VIP process (achieved abortion, incomplete interruption and non-managed abortion) in a selected time period.
- **Characterization of the patient group on the number of failures of VIP process, by year and by month:** identifies the number of patients whose VIP process has not been successfully achieved.
- **Characterization of the patient group based on contraception early in the process by date:** identifies the number of patients who already used contraception in the process, which contraception and which faults are associated over a given period of time.
- **Characterization of the patient group based on the end contraception of VIP process by date:** Identifies the number of patients who have adopted contraception at the end of the process and which contraception was chosen over a given period of time.
- **Characterization of the patient group based on gestational age at the time of taking the mifepristone, by year, month and day:** identifies the number of patients for each of the possible gestational ages, by year, month and day. To this end, it was created an OLAP cube that allows the drill-down and roll-up on the date, providing visualization of the months and days for a given year, as shown in Figure 3.

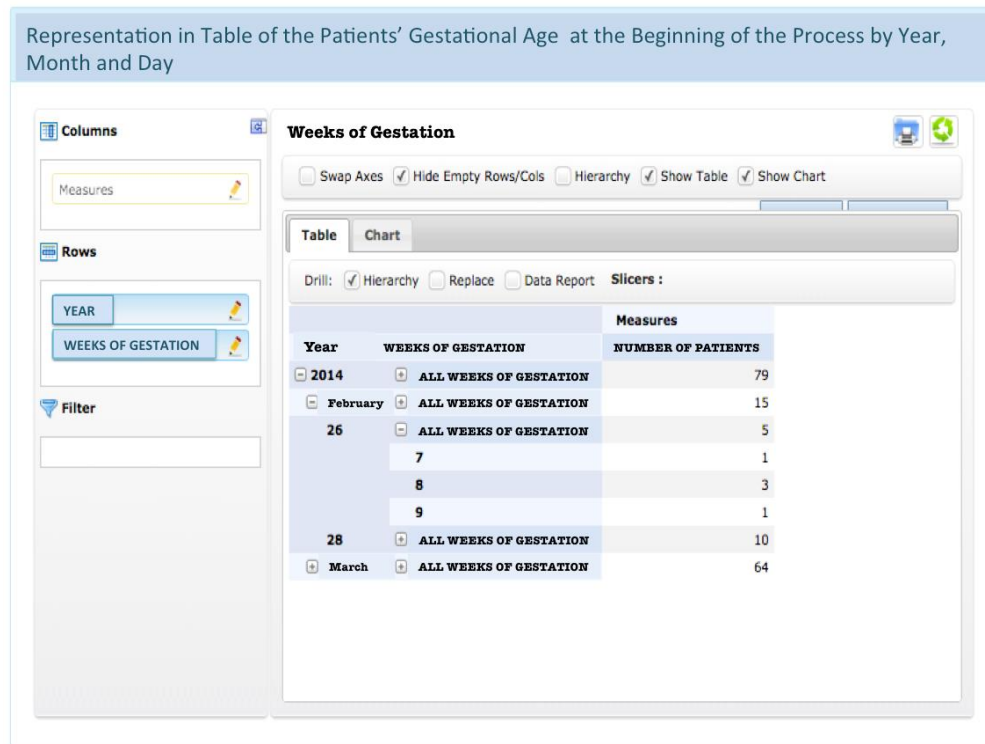


Figure 3. Characterization of the user group based on gestational age at the time of taking the mifepristone, by year, month and day.

- **Characterization of the patient group based on the process achievement local, by date:** identifies the number of patients, based on the achievement local of the process (home or CMIN), by date.
- **Characterization of the patient group by profession and by date:** identifies the number of patients who are unemployed, students and employees, while setting the respective professions in a given period of time. This indicator is shown as a pie chart and two bar graphs, as can be seen in Figure 4.

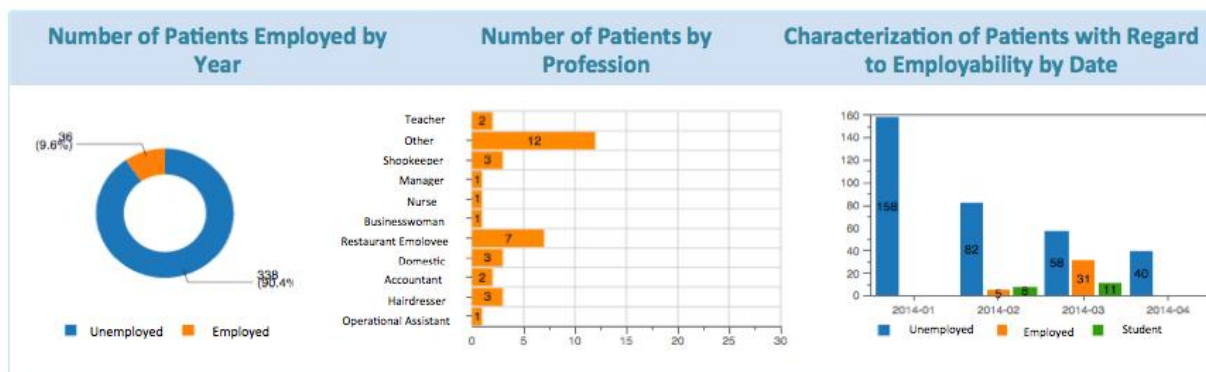


Figure 4. Characterization of the user group by profession and by date.

- **Characterization of the patients group by age and year:** identifies the number of patients by age classes in a given year.

Furthermore, the BI platform in CMIN also incorporates the DM prediction models, discussed above. These models predict where the second dose of the drug should be administered and predict which are the probability of a patient belonging to the group risk. The way that prediction models are represented in BI platform can be seen in Figure 5. For example in the case of group risk a patient with a probability between 60% and 100% (red color) have a higher probability of belonging to the group risk. In the case of this example, the patient probability is 30%.

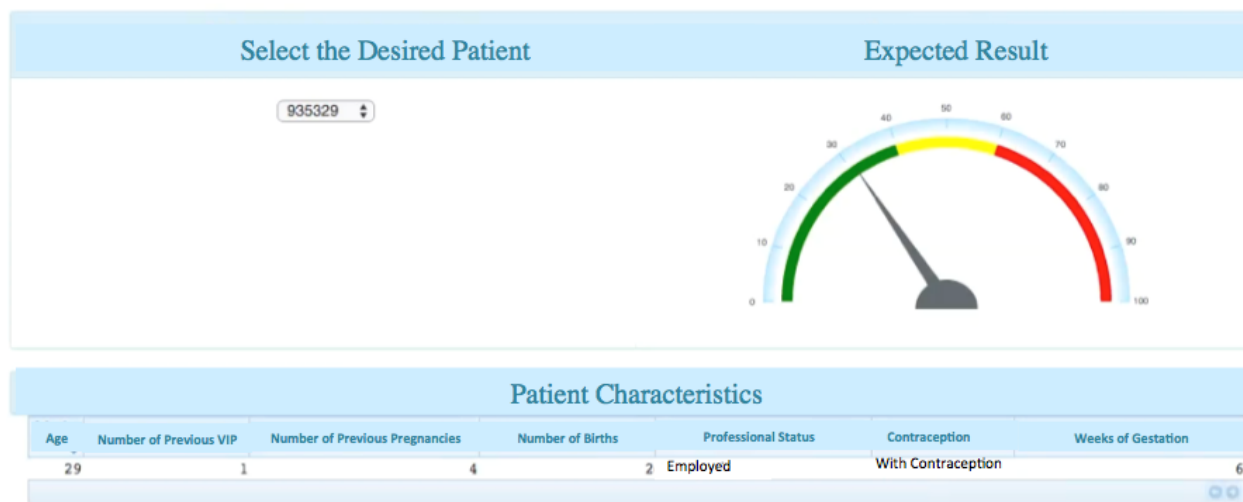


Figure 5. Representation of Prediction Models in BI Platform

## DISCUSSION

Given the DM models developed for each one of the problems discussed, it can be concluded that the results were quite acceptable based on the assessment carried out for the forecast models. For the first problem, it was obtained sensitivity and acuity values of approximately 91% and 87%, respectively, in model 3, resulting from the application of the technique Decision Trees. Thus, it can be considered that the most important factors for the determination of the second dosage of medication administration place is the number of pregnancies and the number of births that a patient had. The second DM question, gave a maximum of approximately 93% in the model 4, in terms of sensitivity metric, using Support Vector Machine technique. Thus, it appears that the most relevant factors in determining the patients belonging to the risk group are the age, the number of VIPs previously performed, the number of pregnancies and the number of births.

In terms of the development of the BI platform to the VIP unit, there are many benefits of integrating BI technology in CMIN. The main advantage is the autonomy and flexibility that the end users acquire with regard to reporting. It also allows a simply and quickly data analysis resulting in an improvement of the decision-making process, a greater operational efficiency and a greater variety of new analytics.

At the scientific level this platform has many advantages. In fact it is a pioneering worldwide project in the VIP area, characterized by obtaining specific indicators to maternity care institutions. Taking into account the results obtained and the accreditation given by health professionals, this platform can provide a starting point for the implementation of improvement measures in the VIP module. Moreover, it may still be useful for the research and development of new modules of interest not only for the VIP module

as well as for other health institutions. An example of another module of interest may be the case with non-progressive pregnancy, since they require only a small adaptation of the current system conditions to some new in this procedure.

The implementation of the BI platform in VIP unit is very valuable, since it allows, for example, the interpretation of existing data records, obtaining indicators reflecting the overall operation of this unit, the acquisition of continuous improvement strategies in care based on the results, and the disclosure of the activities carried out in this area.

The main objective of this BI platform is to support the provision of better health care and meet the needs of pregnant women who wish to undergo the VIP process. With this platform, it is aimed to provide personalized care to the patient, enabling an increase in evidence-based clinical practice and providing new knowledge in real-time.

## **FUTURE RESEARCH DIRECTIONS**

In the future can be explored new techniques and create more DM problems. New variables can be selected and new experiments can be performed with new datasets. In terms of BI platform a set of new indicators will be designed in order to provide personalized care to the patients

It is also expected a gradual improvement platform, responding to the needs of healthcare professionals in CMIN. In the future, platform could also be expanded to other units of interest in CMIN, as well as other maternal and child organizations.

## **CONCLUSION**

The BI platform in CMIN can bring many benefits to this institution, namely, increased autonomy and flexibility of the users with regard to creating reports, faster and easier analysis of the clinical data, support clinical decision-making, among others.

With this chapter, it was also demonstrated that it is possible to obtain DM classification models to whether predict where the VIP patient must carry the administration of the second dose of medication, as for predicting whether a patient belongs to the risk group in VIP process. This study was conducted using real data of VIP processes, collected in CMIN, corresponding to a year of operation, namely the year of 2012. It can be concluded then, using data mining classification techniques and the historic data of the patients of the VIP unit CMIN, it is possible provide the most appropriate location for the realization of the second stage of the VIP process, as well as identify patients with greater ability to belong to the risk group in the VIP process. Thus, it was demonstrated the possibility of developing DM models with good predictive ability in order to provide the pregnant woman better treatment under appropriate conditions. These problems of DM described are intended only for VIP, in particular in CMIN however, they can be applied in other hospitals, which has the same problems and use the same procedures in the VIP process.

In addition, this chapter allowed exploring the importance of integrating BI technology in CMIN. The implementation of this technology provides an analysis of the extracted data, while the majority of clinical data are unstructured.

Finally, it can be concluded that the implementation of the BI platform including DM models in CMIN is an innovative idea, as it is the first time that the BI technologies are applied in the VIP area.

## **ACKNOWLEDGES**

This work has been supported by FCT - Fundação para a Ciência e Tecnologia within the Project Scope UID/CEC/00319/2013.

## REFERENCES

- Abelha, A., Analide, C., Machado, J., Neves, J., Santos, M., & Novais, P. (2007). Ambient Intelligence and Simulation in Health Care Virtual Scenarios. In L. M. Camarinha-Matos, H. Afsarmanesh, P. Novais, & C. Analide (Eds.), *Establishing the Foundation of Collaborative Networks* (pp. 461–468). Boston, MA: Springer US. doi:10.1007/978-0-387-73798-0
- Bonney, W. (2013). Applicability of Business Intelligence in Electronic Health Record. *Procedia - Social and Behavioral Sciences*, 73, 257–262. doi:10.1016/j.sbspro.2013.02.050
- Cabral, A., Abelha, A., Salazar, M., Quintas, C., Portela, F., Machado, J., Neves, J., & Santos, M. F. (2013). *Knowledge acquisition process for intelligent decision support in critical health care*. IGI Global Book. Retrieved from <http://repositorium.sdum.uminho.pt/handle/1822/21710>
- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). The CRISP-DM User Guide. *NCR Systems Engineering Copenhagen*. Brussels: NCR Systems Engineering Copenhagen.
- Cios, K., Pedrycz, W., Swiniarski, R., & Kurgan, L. (2007). *Data Mining. A knowledge Discovery Approach*. Springer.
- El-Sappagh, S. H. A., Hendawi, A. M. A., & El Bastawissy, A. H. (2011). A proposed model for data warehouse ETL processes. *Journal of King Saud University - Computer and Information Sciences*, 23(2), 91–104. doi:10.1016/j.jksuci.2011.05.005
- Fayyad, U., Piatetsky-shapiro, G., & Smyth, P. (1996). From Data Mining to Knowledge Discovery in Databases. *American Association for Artificial Intelligence*, 37–54.
- Gonçalves, J., Portela, F., Santos, M. F., Silva, Á., Machado, J., Abelha, A., & Rua, F. (2013). Real-time Predictive Analytics for Sepsis Level and Therapeutic Plans in Intensive Care Medicine. *International Information Institute*.
- Khodambashi, S. (2013). Business Process Re-engineering Application in Healthcare in a Relation to Health Information Systems. *Procedia Technology*, 9, 949–957. doi:10.1016/j.protcy.2013.12.106
- Kimball, R., & Ross, M. (2002). The Data Warehouse Toolkit. *John Wiley and Sons, Inc.* Retrieved October 09, 2014, from [http://home.elka.pw.edu.pl/~rbzoma/The Data Warehouse Toolkit - The Complete Guide to Dimensional Modeling \(2nd Ed 2002 Wiley\) - Kimball & Ross.pdf](http://home.elka.pw.edu.pl/~rbzoma/The Data Warehouse Toolkit - The Complete Guide to Dimensional Modeling (2nd Ed 2002 Wiley) - Kimball & Ross.pdf)
- Machado, J., Alves, V., Abelha, A., & Neves, J. (2007). Ambient intelligence via multiagent systems in the medical arena. *Engineering Intelligent Systems for Electrical Engineering*

*and Communications* (2002), 15(3), 151–157. Retrieved from <http://www.refdoc.fr/Detailnotice?idarticle=696527>

Mettler, T., & Vimarlund, V. (2009). Understanding Business Intelligence in the Context of Healthcare. *Health Informatics Journal*, 3(15), 254–264.

Palaniappan, S., & Awang, R. (2008). Intelligent Heart Disease Prediction System Using Data Mining Techniques. In *Proceedings of the 2008 IEEE/ACS International Conference on Computer Systems and Applications* (pp. 108–115). Washington, DC, USA: IEEE Computer Society. doi:10.1109/AICCSA.2008.4493524

Peixoto, H., Santos, M., Abelha, A., & Machado, J. (2012). Intelligence in Interoperability with AIDA. In L. Chen, A. Felfernig, J. Liu, & Z. Raś (Eds.), *Lecture Notes in Computer Science, Foundations of Intelligent Systems - 31* (Vol. 7661, pp. 264–273). Springer Berlin Heidelberg. doi:10.1007/978-3-642-34624-8\_31

Pinto, L. F. B. (2009). *Sistemas de Informação e Profissionais de Enfermagem*. Universidade de Trás-os-Montes e Alto Douro.

Portela, F., Vilas-Boas, M., Santos, M. F., Abelha, A., Machado, J., Cabral, A., & Aragão, I. (2010). Electronic Health Records in the Emergency Room. *Computer and Information Science (ICIS), 2010 IEEE/ACIS 9th International Conference on*. doi:10.1109/ICIS.2010.98

Raquel, O., & Oliveira, F. (2012). *Extração de Conhecimento nas Listas de Espera para Consulta e Cirurgia*. Universidade do Minho.

Reinschmidt, J., & Francoise, A. (2000). *Business Intelligence Certification Guide*. IBM.

Steiner, M., Soma, N., & Shimizu, T. (2006). Abordagem de um problema médico por meio do processo de KDD com ênfase à análise exploratória dos dados. *Gest Prod*. Retrieved from <http://www.scielo.br/pdf/gp/v13n2/31177.pdf>

Valente, C., Cristina, T., Rosário, F., & Alcina, B. (2012). *Acompanhamento de enfermagem na interrupção da gravidez por opção da mulher ( I.G.O.)*. Porto.

Abelha, A., Pereira, E., Brandão, A., Portela, F., Santos, M. F., Machado, J., et al. (2015). Improving Quality of Services in Maternity Care Triage System. *International Journal of E-Health and Medical Communications (IJEHMC)*, 6(2), 10-26.

Brandão, A., Pereira, E., Portela, F., Santos, M., Abelha, A., & Machado, J. (2014). *Real-time Business Intelligence platform to maternity care*.

Brandão, A., Pereira, E., Portela, F., Santos, M. F., Abelha, A., & Machado, J. (2014). Managing Voluntary Interruption of Pregnancy Using Data Mining. *Procedia Technology*, 16, 1297-1306.

- Cardoso, L., Marins, F., Portela, F., Abelha, A., & Machado, J. (2014). Healthcare interoperability through intelligent agent technology. *Procedia Technology*, 16, 1334-1341.
- Cardoso, L., Marins, F., Portela, F., Santos, M., Abelha, A., & Machado, J. (2014). The Next Generation of Interoperability Agents in Healthcare. *International journal of environmental research and public health*, 11(5), 5349-5371.
- Filipe Portela, Jorge Aguiar, Manuel Filipe Santos, Álvaro Silva, & Fernando Rua. (2013). Pervasive Intelligent Decision Support System - Technology Acceptance in Intensive Care Units. In Springer (Ed.), *Advances in Intelligent Systems and Computing*: Springer.
- Machado, J. M., Abelha, A., Santos, M., Portela, F., Pereira, E., & Brandão, A. (2015). Predicting the risk associated to pregnancy using data mining, ICAART 2015.
- Marins, F., Cardoso, L., Portela, F., Santos, M. F., Abelha, A., & Machado, J. (2014). Improving High Availability and Reliability of Health Interoperability Systems. In *New Perspectives in Information Systems and Technologies, Volume 2* (pp. 207-216): Springer.
- Pereira, E., Brandão, A., Portela, C. F., Santos, M. F., Machado, J., & Abelha, A. (2014). *Business intelligence in maternity care*, IDEAS 2014.
- Portela, C. F., Santos, M. F., Silva, A., Machado, J., & Abelha, A. (2011). Enabling a Pervasive Approach for Intelligent Decision Support in Critical Health Care. In M. M. CruzCunha, J. Varajao, P. Powell & R. Martinho (Eds.), *Enterprise Information Systems, Pt 3* (Vol. 221, pp. 233-243).
- Portela, F., Gago, P., Santos, M. F., Silva, A., Rua, F., Machado, J., et al. (2011). Knowledge Discovery for Pervasive and Real-Time Intelligent Decision Support in Intensive Care Medicine. In S.-A. t. publication (Ed.). In *KMIS 2011- International Conference on Knowledge Management and Information Sharing*, pp. 12). Paris , France:(Springer - Accepted for publication).
- Portela, F., Santos, M. F., Silva, Á., Rua, F., Abelha, A., & Machado, J. (2013). Adoption of Pervasive Intelligent Information Systems in Intensive Medicine. *Procedia Technology*, 9, 1022-1032.
- Portela, F., Santos, M. F., & Vilas-Boas, M. (2012). A Pervasive Approach to a Real-Time Intelligent Decision Support System in Intensive Medicine. In Springer (Ed.), *Communications in Computer and Information Science* (Vol. 0272, p. 14).